Studies that find differences between the quality of comprehension of interlocutors and of overhearers (e.g., Schober & Clark 1989; Wilkes-Gibbs & Clark 1992) might be taken as evidence to support the idea of radically different cognitive processes in interactive discourse. They need not be. In such studies interlocutors had opportunities for feedback and repair that the overhearers lacked. Because different people will misunderstand different things, those who can ask for clarification will receive feedback that is relevant to them and consequently might understand better. Therefore any difference between such noninteractive and interactive comprehension could be fully attributable to strategic, effortful feedback but not necessarily to automatic alignment. In fact, Barr and Keysar (2002) found that even when such feedback is removed, listeners who believed themselves to be overhearers automatically aligned their semantic representations with the speaker's to the same degree as listeners who believed themselves to be addressees.

In closing, far from qualitatively changing the nature of processing, it is likely that dialogue provides a radically different context in which the same processes operate. The context includes an interlocutor and mechanisms for feedback and interactive repair. For us there is no question that it is important to study conversation in vivo, but it remains to be seen whether this would reveal automatic processes that are truly unique to dialogue.

NOTE

1. In their seminal work on common ground, Clark and Marshall (1981) clearly make the case that common ground is a form of metaknowledge that is conceptually distinct from shared knowledge. What P&G are referring to by "implicit common ground" is really just shared knowledge, not common ground, because interlocutors need not represent the fact that their representations are shared. Such usage is certain to contribute to the legacy of confusion that has plagued discussions of mutual knowledge and common ground (see Keysar 1997 and Lee 2001 for discussion).

Full alignment of some but not all representations in dialogue

Holly P. Branigan

Department of Psychology, University of Edinburgh, Edinburgh EH8 9JZ, United Kingdom. **holly@cogsci.ed.ac.uk**
http://www.psy.ed.ac.uk/Staff/ra/hbranigan/publications.html

Abstract: I argue that alignment of linguistic representations and situation models in dialogue are qualitatively distinct. By virtue of the isomorphy between interlocutors' linguistic representations, interlocutors align their linguistic representations fully. However, evidence about situation models is indirect and mediated through language, with the result that alignment of situation models is only partial.

Pickering & Garrod (P&G) provide a plausible and very welcome account of language processing in dialogue. Their account assigns a central importance to the notion of alignment. Here I consider the nature of alignment and, in particular, whether alignment of strictly linguistic representations and alignment of situation models are qualitatively distinct.

P&G suggest that successful dialogue arises from the alignment of representations between interlocutors, and particularly alignment of situation models. Alignment is defined as interlocutors having the same representation at a particular level of structure. It is uncontroversial that language is represented in the same way in different speakers. Linguists and psycholinguists assume a common competence grammar in adult speakers (Chomsky 1965), even though they may differ as to whether this grammar is innate (e.g., Pinker 1989) or to some extent constructed through experience (e.g., Tomasello 2000). Crucially, adult speakers' internalised knowledge of the syntax, lexicon, and morpho-phonology of a language is held to be identical, such that there is an isomorphic mapping from any one speaker's internalised representation of the language to any other's. In a dialogue, then, interlocutors necessarily make use of identical representations in producing their utterances.

Under P&G's maximally parsimonious assumption of parity of representations, interlocutors also necessarily draw upon identical representations in both producing and comprehending utterances. Note also that speakers' utterances provide direct linguistic evidence to the listener. So when a listener hears an utterance, he receives direct evidence (except in cases of mishearing or unresolved structural ambiguities) about the syntactic, lexical, and morpho-phonological representations that the speaker has employed. If a speaker produces an utterance like *I am in row two,* for example, the listener has direct evidence that she has used the words *I, am,* and so on (and their relevant inflectional markings), that she has used a pronoun and a verb and so on, and that she has used a noun phrase, a verb phrase, a prepositional phrase, and so on. Taken together, the combination of isomorphy of representations and direct evidence strongly supports P&G's contention (summarised in their Fig. 2) that linguistic representations used by interlocutors in dialogue act directly upon one another, and that, in a very real sense, when we talk about interlocutors having aligned linguistic representations, we mean that those representations are identical. In summary, P&G's arguments for full alignment at linguistic levels of representation seem well founded.

But are situation models aligned in the same way as linguistic representations? It is unclear that this is the case. In P&G's model, interlocutors' situation models act directly upon one another (see the authors' Fig. 2), in the same way as syntactic, lexical, and morpho-phonological representations do, and alignment of situation models is taken as critical for successful communication. But situation models differ qualitatively from strictly linguistic representations. A speaker's utterances do not give direct evidence of the situation model that the speaker holds, only indirect evidence encoded in linguistic representations, from which the listener has to infer the speaker's situation model. So whereas an utterance like *I am in row two* gives direct evidence about the speaker's syntactic, lexical, and morpho-phonological representations, it gives only indirect evidence about the speaker's situation model. The listener must construct a situation model based upon his or her interpretation of the speaker's meaning – which may or may not be correct. Of course, as P&G note, misunderstandings may come to light, and interlocutors may initiate repairs to bring about situation models that are aligned in the relevant aspects. But as they also note, some misunderstandings may not be repaired. In fact, it seems likely that interlocutors quite frequently have situation models that are misaligned in major respects. Communication will be (apparently) successful as long as the misalignment is not apparent to the interlocutors. To take P&G's example of interlocutors using *John* to refer to different people, it is quite possible for them to have a mutually satisfying dialogue concerning this person without ever realising that they are discussing different people; unless one of them says something that is inconsistent with the other's knowledge, they can successfully (for their purposes) complete a dialogue with quite radically different situation models. Equally, a doctor and a patient may have a dialogue concerning the patient's *chronic back problem* that appears to be successful, in that they are both satisfied that they understand each other well; yet their situation models may differ considerably because of unresolved (and unapparent) differences in their interpretation of *chronic.* Situation models need only be aligned sufficiently for the current communicative goal to be (apparently) met.

So it seems that alignment of situation models and alignment of linguistic representations are quite different. With linguistic representations, interlocutors genuinely employ aligned (i.e., identical) representations that act directly upon one another; whereas, because evidence for situation models is only indirect, interlocu-

https://doi.org/10.1017/S0140525X04230059 Published online by Cambridge University Press

tors rarely if ever have *identical* models. Instead, they have partially aligned models that may differ in many – sometimes important – respects. And because evidence for situation models is mediated through language, it seems highly unlikely that they can act directly upon one another (contra P&G's Fig. 2).

One interesting result of the distinction between alignment of linguistic and situation models is that alignment of linguistic representations may sometimes lead to misaligned situation models. Garrod and Clark (1993) found that young children had a tendency to use the same words to describe a maze – that is, showed lexical alignment – even when their situation models were quite different. Similarly, in the case of the doctor-and-patient scenario, one speaker's use of the term *chronic* may well reinforce the other's use of the same term, leading to more misunderstanding than if a different term were used. In both examples, full alignment at the linguistic level misleads interlocutors into believing that they also have alignment at the level of situation models.

ACKNOWLEDGMENTS

This research was supported by a British Academy Postdoctoral Fellowship and by ESRC grant R000239363.

Two steps forward, one step back: Partner-specific effects in a psychology of dialogue

Susan E. Brennan and Charles A. Metzing

Department of Psychology, State University of New York, Stony Brook, NY 11794-2500. **susan.brennan@sunysb.edu cmetzing@ic.sunysb.edu**
http://www.psychology.sunysb.edu/psychology/personnel/Susan.htm

Abstract: Pickering & Garrod's (P&G's) call to study language processing in dialogue context is an appealing one. Their interactive alignment model is ambitious, aiming to explain the converging behavior of dialogue partners via both intra- and interpersonal priming. However, they ignore the flexible, partner-specific processing demonstrated by some recent dialogue studies. We discuss implications of these data.

In human language processing, the whole is greater than the sum of the parts; therefore, those who study the language processing system in dialogue contexts are poised to make different sorts of discoveries than those who study the parts working alone. Pickering & Garrod (P&G) present a convincing argument that psycholinguists should pay attention to dialogue. In fields such as artificial intelligence and human-computer interaction, where the goal is often to build a fully working dialogue partner, many will find this a worthy enterprise as well. After presenting evidence for phonological, lexical, and syntactic convergence between dialogue partners and for representations shared between comprehension and production, P&G make a strong claim that is far *less* convincing: "normal conversation does not routinely require modeling the interlocutor's mind" (sect. 4.4, para. 4). They support this position with evidence from studies that fail to meet the very standards they seek to advance, while ignoring evidence that complicates matters for their interactive alignment model. Thus, their position on the importance of studying language in dialogue does not go far enough.

This position assumes that interlocutors achieve aligned mental representations without having to track anything specific about each other's knowledge because both have evolved with the same cognitive architecture; what is easiest for speakers is easiest for addressees (Brown & Dell 1987). It further assumes that there is no need to track common ground, as interlocutors each use their own memory of the conversation as a proxy. By this argument, what appears to be partner-specific or "audience design" is actually inflexible and unavoidable, at least in the earliest moments of processing. P&G propose a two-stage model (similar to that of Horton & Keysar 1996), arguing that interlocutors "do not routinely take common ground into account during initial processing . . . full common ground is only used when simpler mechanisms are ineffective" (sect. 4.1, last para.). This (circular) view relegates any aspect of production or interpretation that displays flexibility or sensitivity to an interlocutor's needs (as distinct from one's own) to the status of a relatively late adjustment, managed as a kind of repair or pragmatic garden path.

Granted, it is difficult to design a good experiment on audience design. A good experiment must distinguish one interlocutor's perspective from another's, avoid confounding individual perspectives with common ground (Keysar 1997), and allow interlocutors to interact naturally or contingently (Schober & Brennan 2003). But we are surprised that studies succeeding in all this (and finding partner-specific effects early in processing, e.g., Hanna et al. 2003; Nadig & Sedivy 2002) are dismissed by P&G: "their task was repetitive and involved a small number of items, and listeners were given explicit information about the discrepancies in knowledge" (target article, sect. 4.2, para. 3). Then follows a very broad claim: "Under such circumstances, it is not surprising that listeners develop strategies that may invoke full common ground. During natural dialogue, we predict that such strategies will not normally be used."

Paradoxically, evidence to support this position comes mainly from studies that did not allow any potential for interaction. These include Brown and Dell (1987), Ferreira and Dell (2000), Horton and Keysar (1996), and others in which partners did not interact naturally or provide contingent feedback. Sometimes this matters; for example, Brown and Dell (1987) concluded that speakers did not take addressees' specific needs into account when retelling stories; but their addressees had no needs (they were confederates who knew the stories better than the speakers did). When we ran a similar study using spontaneously interacting speakers and addressees (Lockridge & Brennan 2002), speakers' early syntactic choices indeed showed sensitivity to addressees' needs.

There is additional good evidence of rapid, partner-specific effects from the comprehension side. Hanna and Tanenhaus (2004) asked addressees to follow a (confederate) speaker's directions in a cooking task (e.g., *Hand me the cake mix*); the addressees' eye fixations showed that they restricted candidate referents for ambiguous expressions (e.g., when two cake mixes were present) depending on what the speaker was holding and what she could not reach; they did this from the earliest moments of processing.

And we have demonstrated that addressees interpret the same utterance differently when it is spoken by different speakers with whom the addressees have different dialogue histories (Metzing & Brennan 2001; 2003). In our experiment, addressees were instructed by (confederate) speakers to reposition objects among a relatively large set; they did this several times, evolving shared perspectives and terms for critical objects (e.g., *the shiny cylinder*). Then the speaker left the room and either returned or else a new confederate speaker entered. In the final trial, the new or old speaker used either the familiar term or a new, equally good term (e.g., *the silver pipe*) for the same critical object (amid many other references that did not use different terms). Addressees gazed immediately at the object when either speaker used the old term. However, when the old speaker used a new term (inexplicably breaking a conceptual pact), addressees experienced interference, delaying gazing at the target object. There was no such delay when the new speaker used the new term (in fact, resolving this was just as fast as the old term spoken by the new speaker). This partner-specific interference suggests that the pragmatic force of breaking a conceptual pact has impact immediately, rather than just as a late adjustment or repair.

Such immediate effects provide evidence of impressive agility and potential for partner-specific processing in the language processing system, which the interactive alignment proposal fails to address. Pragmatic and partner-specific knowledge is implemented by basic mechanisms of memory and does not rely on special processes or exhaustive partner models. Audience design – truly partner-specific processing – can occur immediately and effortlessly as well as more slowly and deliberately, depending on

https://doi.org/10.1017/S0140525X04230059 Published online by Cambridge University Press